UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/063,735	05/08/2002	Dan L. Eaton	P3230R1C001-168	2793

30313 7590 10/06/2004

KNOBBE, MARTENS, OLSON & BEAR, LLP
2040 MAIN STREET
IRVINE, CA 92614

EXAMINER	
SEHARASEYON, JEGATHEESAN	
ART UNIT	**PAPER NUMBER**
1647	

DATE MAILED: 10/06/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/063,735	EATON ET AL.	
	Examiner	Art Unit	
	Jegatheesan Seharaseyon	1647	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *10 September 2002*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-20* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *11 and 12* is/are allowed.

6)☒ Claim(s) *1-10 and 13-20* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *9/17/2002*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☒ Other: *See Continuation Sheet*.

Continuation of Attachment(s) 6). Other: Notice to comply Appendix A & B.

DETAILED ACTION

1. Applicant's preliminary amendment filed on 10 September 2002 is acknowledged and entered. Claims 1-20 are pending and under consideration. The claims are drawn to the nucleotide encoding protein designated PRO1774, also identified as encoded by DNA77626-1705 and ATCC accession number 203536, shown in Figures 127 (nucleic acid) and 128 (protein).

Specification

2. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

3. This application contains sequence disclosures that are encompassed by the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR 1.821(a)(1) and (a)(2). However, this application fails to comply with the requirements of 37 CFR 1.821 through 1.825 for the reason(s) set forth on the attached Notice To Comply With Requirements For Patent Applications Containing Nucleotide Sequence And/Or Amino Acid Sequence Disclosures. Applicant must comply with the requirements of the sequence rules (37 CFR 1.821 - 1.825). **Applicant is required to provide a paper copy of the CRF in response to the Office Action.**

Information Disclosure Statement

4. The information disclosure statement, filed 9/17/2002, has been considered. The BLAST results demonstrate that applicants are aware of nucleic acids with identity/homology to the one claimed herein. However, as the BLAST results do not

give sufficient identifying information, the Examiner cannot determine if said sequences constitute prior art.

Priority Determination

5. Since no utility is disclosed in the priority applications, they aren't enabling under 35 USC 112, as required under 119(e), thus no priority is granted. Accordingly, priority under 35 U.S.C. 120 is set at the instant filing date, 5/8/02.

Should the applicant disagree with the examiner's factual determination above, it is incumbent upon the applicant to provide the serial number and specific page number(s) of any parent application filed prior to the date recited above which specifically supports the particular claim limitation for each and every claim limitation in all the pending claims which applicant considers to have been in possession of, and fully enabled for, prior to that date.

6. The utility for the instant invention is based on the sequence identity of SEQ ID NO: 128 to that of human alcohol dehydrogenase described by Meyers et al. (AAB84364, WO 01/44446, Pub. Date 06/01). The role of ADH activity in colorectal cancer is described in page: 4.

Claim Rejections - 35 USC § 112

7. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1-6, 8-10 and 14-20 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

7a. The protein identified as PRO1774 (SEQ ID NO: 128) is not disclosed as being expressed on a cell surface. Accordingly, the limitation that the claimed protein comprises an "extracellular domain" (for example see claims 1, 6 and 14 parts (c) and (d)) is indefinite, as the art does not recognize soluble proteins as having such domains. Further, if the protein had an extracellular domain, the recitation of "the extracellular domain", "lacking its associated signal sequence" (claim 1, 6 and 14, part (d), for example) is indefinite as a signal sequence is not generally considered to be part of an extracellular domain, as signal sequences are cleaved from said domains in the process of secretion from the cell. Claims 2-5, 8-10 and 15-20 are rejected insofar as they are depended on rejected claims 1, 6 and 14.

7b. Claims that recite that the claimed polynucleotide "hybridizes to" another sequence, such as claim 14, are indefinite as there is no limiting definition of such in the specification, and the metes and bounds of that which will hybridize are dependent upon the conditions under which the hybridization is performed. As the metes and bounds of what will hybridize to a given sequence are entirely dependent upon the conditions of hybridization and washing, the metes and bounds of the claims cannot be determined. With respect to claim 15, although the further limitation that the hybridization conditions are "stringent" is introduced, the term "stringent conditions" is also a relative term, and

the metes and bounds of the claim cannot be determined. Claim 15 is rejected insofar as it is depended on rejected claim 14.

8. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

8a.Claims 1-5 and 14-20 are rejected under 35 U.S.C. 112, first paragraph, because the specification, while being enabled only for the polynucleotide of SEQ ID NO: 127 or fragments of such that are usable as hybridization probes and are not enabled for polynucleotides 80, 85, 90, 95 or 99% identical to such, nor which encode a protein 80, 85, 90, 95 or 99% identical to the protein of SEQ ID NO: 128, nor polynucleotides which hybridize to any of the above because there is no structural or functional information provided in the specification. Therefore, the specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the invention commensurate in scope with these claims.

The factors considered when determining if the disclosure satisfies the enablement requirement and whether any necessary experimentation is "undue" include, but are not limited to:

1) nature of the invention, 2) state of the prior art, 3) relative skill of those in the art, 4) level of predictability in the art, 5) existence of working examples, 6) breadth of claims, 7) amount of direction or guidance by the inventor, and 8) quantity of experimentation needed to make or use the invention. In re *Wands*, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988).

The claims are directed to isolated polynucleotides having at least 80% identity to a SEQ ID NO: 127 or that encode the protein of SEQ ID NO: 128 with or without its signal peptide, or which encode the extracellular domain of SEQ ID NO: 128 with or without its signal peptide, or polynucleotides at least 80% identical to such encoding polynucleotides. Dependent claims are directed to polynucleotides that hybridize to the above sequences, vectors and host cells comprising the isolated polynucleotides. In the instant application, there is insufficient guidance regarding how to make PRO1774 polynucleotides variants recited in the claims.

The specification also is not enabling for the breadth of claims to polynucleotide molecules that hybridize to the disclosed sequences. It is noted that claims that recite hybridization language fail to provide adequate guidance, and do not recite that the polynucleotide encodes a protein, much less one having a specifically disclosed activity. First of all, it is pointed out that the term "hybridize" or "hybridization" generically refers to a process in which a strand of polynucleotide joins or matches up with a complementary strand through the process of base pairing, wherein the process is basically used to locate or identify DNAs encoding specific proteins. It is well established in the art that 15-20 bases have been considered sufficient to achieve this process. The breadth of the claims includes polynucleotides of as little as 10 nucleotides. With these points in mind, it is the Examiner's position that giving the claims their broadest reasonable interpretation, this language reads on an infinite number of possible DNA sequences for which there is not sufficient enablement without

undue experimentation because of the breath of claims, the lack of guidance provided and the quantity of experimentation needed to make or use the invention.

With respect to the hybridization use, as discussed above in paragraph 6 the invention lacks utility and thus lacks enablement. Even if utility were established, the enablement is commensurate in scope only with claims to polynucleotides that are fragments of SEQ ID NO: 127, said fragments of sufficient length to be used as hybridization probes or primers. However, enablement is *not* commensurate in scope with fragments of polynucleotides that differ from SEQ ID NO: 127 due to codon degeneracy, as it is not recognized in the art to use such sequences that are degenerate for such detection or synthesis, and the specification provides no guidance as to how or why to make such degenerate probes or primers. The specification also is not enabling for the breadth of claims to polynucleotide molecules that hybridize to the disclosed sequences because of the quantity of experimentation needed and the lack of guidance provided by the inventor.

The examples provided in the specification do not provide working examples of different DNA sequences that would enable a representative number of the above discussed DNA sequences with assurances that they can be used as probes or primers for the purpose of amplifying or detecting the PRO1774 gene. The mere recitation of this term, and the definitions provided do not serve as sufficient guidance to enable the breadth of the claims for the various DNA sequences claimed. See Ex parte *Forman*, 230 USPQ 546. Since the first paragraph of the statute under 35 U.S.C. 112 requires that there must be an enabling disclosure to support the breadth of the Claims, a review

of the specification confirms that the scope of the various DNA sequences that are discussed above have not been enabled. There is but a single polynucleotide disclosed with reference to PRO1774, SEQ ID NO: 127. In the absence of working examples, breadth of claims and sufficient guidance, it would require undue experimentation to enable a commensurate number of the sequences that are encompassed by the claims.

Since the claimed polynucleotides are described at least in part in terms of the protein that might be encoded, the scope of the protein itself must be considered: The specification asserts that PRO1774 is an unspecified secreted and transmembrane polypeptide. However, this family of proteins does not possess a common utility, but rather the proteins that can be broadly classified and have different activities, that confer different uses on them. Accordingly, the mere identification of a protein as belonging to a family, while indicative of evolutionary relatedness, is not indicative of function, nor by extension, of utility. The structure of the putative PRO1774 peptide is briefly discussed in Figure 128, as having a putative signal sequence, corresponding to amino acids 1-17. It also describes N-myristoylation sites, corresponding to about amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115 and 199-205. In addition it also describes short-chain alcohol dehydrogenase sequences at amino acids 30-42 and 104-114.

The problem of predicting protein structure from sequence data and in turn utilizing predicted structural determinations to ascertain functional aspects of the protein is extremely complex. While it is known that many amino acid substitutions are generally possible in any given protein, the positions within the protein's sequence where such amino acid substitutions can be made with a reasonable expectation of

success are limited. Certain positions in the sequence are critical to the protein's structure/function relationship, e.g. such as various sites or regions directly involved in binding, activity and in providing the correct three-dimensional spatial orientation of binding and active sites. However, Applicant has provided little or no guidance beyond the mere presentation of sequence data to enable one of ordinary skill in the art to determine, without undue experimentation, the positions in the protein which are tolerant to change (e.g. such as by amino acid substitutions or deletions), and the nature and extent of changes that can be made in these positions. Therefore, undue experimentation would be required of the skilled artisan to make and/or use the claimed invention in its full scope, i.e. all the polynucleotides with the various percent identities.

8b. Claims 1-5 and 14-20 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

To provide evidence of possession of a claimed genus, the specification must provide sufficient distinguishing identifying characteristics of the genus. The factors to be considered include disclosure of compete or partial structure, physical and/or chemical properties, functional characteristics, structure/function correlation, methods of making the claimed product, or any combination thereof.

The claims are drawn to polynucleotides having at least 80%, 85%, 95% or 99% sequence identity with a particular disclosed sequence, or that merely hybridize to a

disclosed sequence. The claims do not require that the claimed polynucleotide encode a particular protein, nor that any protein encoded thereby possess any particular biological activity, nor any particular conserved structure, or other disclosed distinguishing feature. The specification teaches that PRO1774 has (unspecified) homology to secreted and transmembrane polypeptide. The structure of the putative PRO1774 peptide is briefly discussed in Figure 128, as having a putative signal sequence, corresponding to amino acids 1-17. It also describes N-myristoylation sites, corresponding to about amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115 and 199-205. In addition it also describes short-chain alcohol dehydrogenase sequences at amino acids 30-42 and 104-114. However, there is no functional characteristic associated with these motifs, hence the mere observation that they exist is not probative of function or utility. Further, there is no disclosure that the protein is expected to be a transmembrane protein, nor of any extracellular domain. In this case, the only factor present in the claim is a partial structure in the form of a recitation of percent identity. There is not even identification of any particular portion of the structure that must be conserved. Accordingly, in the absence of sufficient recitation of distinguishing identifying characteristics, the specification does not provide adequate written description of the claimed genus.

Vas-Cath Inc. v. Mahurkar, 19 USPQ2d 1111, clearly states that "applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention*. The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed.*" (See page 1117.) The

specification does not "clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed." (See *Vas-Cath* at page 1116). As discussed above, the skilled artisan cannot envision the detailed chemical structure of the encompassed genus of polypeptides, and therefore conception is not achieved until reduction to practice has occurred, regardless of the complexity or simplicity of the method of isolation. Adequate written description requires more than a mere statement that it is part of the invention and reference to a potential method of isolating it. The compound itself is required. See *Fiers v. Revel*, 25 USPQ2d 1601 at 1606 (CAFC 1993) and *Amgen Inc. v. Chugai Pharmaceutical Co. Ltd.*, 18 USPQ2d 1616.

One cannot describe what one has not conceived. See *Fiddes v. Baird*, 30 USPQ2d 1481 at 1483. In *Fiddes*, claims directed to mammalian FGF's were found to be unpatentable due to lack of written description for that broad class. The specification provided only the human sequence.

Therefore, polynucleotides comprising the sequence set forth in SEQ ID NO: 127 or encoding the protein of SEQ ID NO: 128, or fragments thereof sufficiently long to be used as hybridization probes but not the full breadth of the claims meet the written description provision of 35 U.S.C. 112, first paragraph. Applicant is reminded that *Vas-Cath* makes clear that the written description provision of 35 U.S.C. 112 is severable from its enablement provision (see page 1115).

Claim Rejections - 35 USC § 102

9. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless :

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

9a. Claims 1-10 and 13-20 are rejected under 35 U.S.C. 102(b) as being anticipated by Meyers et al. (AAB84364, WO 01/44446, Pub. Date 06/01).

Meyers et al. discloses nucleotides encoding the amino acid sequence of SEQ ID NO: 128 of the instant invention (Appendix A). In addition it disclose nucleotides that are capable of hybridizing to nucleotides encoding polypeptide of SEQ ID NO: 128 of the instant invention. Further, Meyers et al. have described the expression of nucleotides containing vectors with promoter sequences in bacterial hosts (pages 59, 83-85). With respect to the limitation of "lacking its associated signal peptide" in claims 8 and 10 as Meyers et al. teaches recombinant expression of the said polypeptide, the cDNA would produce the polypeptide identical to the present SEQ ID NO: 128, but lacking its associated signal peptide when transfected into the host cell. Thus, meeting the limitations of claims 1-10 and 13-20. Therefore, claims 1-10 and 13-20 are rejected as being anticipated by Meyers et al. (AAB84364, WO 01/44446, Pub. Date 06/01).

9b. Claims 14-16 are rejected under 35 U.S.C. 102(b) as being anticipated by Birren et al. (AC003042/c, Pub. Date 07/98).

Birren et al. disclose (Appendix B) nucleotides that are capable of hybridizing to nucleotides encoding polypeptide of SEQ ID NO: 128 of the instant invention.

10. Claims 11 and 12 are allowable.

Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jegatheesan Seharaseyon whose telephone number is 571-272-0892. The examiner can normally be reached on M-F: 8:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Brenda Brumback can be reached on 571-272-0961. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

JS 09/04

Brenda Brumback
BRENDA BRUMBACK
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 1600

Notice to Comply	Application No. 10/063,735	Applicant(s) T. Eaton et al.
	Examiner J. Seharaseyon	Art Unit 1647

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

Applicant must file the items indicated below within the time period set the Office action to which the Notice is attached to avoid abandonment under 35 U.S.C. § 133 (extensions of time may be obtained under the provisions of 37 CFR 1.136(a)).

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 C.F.R. 1.821 - 1.825 for the following reason(s):

☐ 1. This application clearly fails to comply with the requirements of 37 C.F.R. 1.821-1.825. Applicant's attention is directed to the final rulemaking notice published at 55 FR 18230 (May 1, 1990), and 1114 OG 29 (May 15, 1990). If the effective filing date is on or after July 1, 1998, see the final rulemaking notice published at 63 FR 29620 (June 1, 1998) and 1211 OG 82 (June 23, 1998).

☑ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 C.F.R. 1.821(c).

☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 C.F.R. 1.821(e).

☐ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 C.F.R. 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing."

☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A Substitute computer readable form must be submitted as required by 37 C.F.R. 1.825(d).

☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable from of the "Sequence Listing" as required by 37 C.F.R. 1.821(e).

☐ 7. Other:

Applicant Must Provide:

☐ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing".

☒ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification.

☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 C.F.R. 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d).

For questions regarding compliance to these requirements, please contact:

For Rules Interpretation, call (703) 308-4216 or (703) 308-2923
For CRF Submission Help, call (703) 308-4212 or 308-2923
PatentIn Software Program Support
Technical Assistance.....703-287-0200
To Purchase PatentIn Software.....703-306-2600

PLEASE RETURN A COPY OF THIS NOTICE WITH YOUR REPLY

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RQELREAQTHIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVL 240

Qy        241 STPAHIQIGDIQMRPTEQVT 260
                ||||||||||||||||||||
Db        241 STPAHIQIGDIQMRPTEQVT 260
```

RESULT 5
AAU76220
ID AAU76220 standard; protein; 260 AA.
XX
AC AAU76220;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human 21620 alcohol dehydrogenase (ADH) protein.
XX
KW Alcohol dehydrogenase; ADH; human; cytostatic; antiinflammatory;
KW cerebroprotective; anti-HIV; immunomodulator; hepatotropic; metastases;
KW pulmonary congestion; Meckel diverticulum; splenic infarction;
KW idiopathic inflammatory bowel disease; jaundice; cholestasis;
KW endometriosis; cerebral oedema; AIDS; leukopaenia; splenomegaly;
KW acquired immmune deficiency disease; lupus erythematosus; dermatitis;
KW lung disease; adult respiratory distress syndrome; skin disease;
KW bronchitis; sarcoidosis; pneumothorax; colon disorder; colitis;
KW Crohn's disease; liver disorder; hepatitis; cirrhosis; brain disorder;
KW meningitis; Alzheimer's disease; Huntington's disease; atherosclerosis;
KW ischaemia; chromosome 17 (17q12-21); 21620; enzyme.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Region   166. .176
FT   /note= "Short chain ADH family signature"
XX
PN US2002010946-A1.
XX
PD 24-JAN-2002.
XX
PF 28-FEB-2001; 2001US-00796089.
XX
PR 15-DEC-1999; 99US-00464039.
PR 15-DEC-2000; 2000WO-US033873.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2002-179233/23.
DR N-PSDB; ABK15172.
XX
PT New human alcohol dehydrogenase (ADH) polynucleotides and polypeptides,
PT useful as targets for diagnosing or treating ADH-related or ADH-mediated
PT disorders, e.g. malignant breast metastases, enema or leukopenia.
XX
PS Claim 9; Fig 1A; 80pp; English.
XX
CC This invention relates to the cDNA and protein sequences of 5 novel human
CC alcohol dehydrogenase molecules. The ADH polynucleotides and polypeptides
CC are useful as targets for diagnosing or treating ADH- related or ADH-
CC mediated disorders, e.g. malignant breast, liver, colon or liver
CC metastases, pulmonary congestion or enema, Meckel diverticulum,
CC idiopathic inflammatory bowel disease, jaundice and cholestasis,
CC endometriosis, cerebral oedema, AIDS, or leukopaenia. The sequences may
CC also be used for treating other diseases or disorders such as spleen
CC disorders (splenomegaly, splenic infarction), lung diseases (adult
CC respiratory distress syndrome, bronchitis, sarcoidosis, pneumothorax),
CC colon disorders (colitis, Crohn's disease), liver disorders (hepatitis,
CC cirrhosis), brain disorders (meningitis, Alzheimer's disease,
CC Huntington's disease), heart and blood vessel disorders (atherosclerosis,
CC ischaemia), skin diseases (lupus erythematosus, dermatitis) and many
CC other diseases listed in the specification. The polynucleotides and
CC polypeptides are also useful in screening methods to identify agonists
CC and antagonists for diagnosis or treatment. In particular, the
CC polypeptides and polynucleotides are useful in drug screening assays in
CC cell-based assays or cell-free systems, as well as for biological assays
CC related to ADHs. The ADH polypeptides are also useful for producing
CC antibodies specific for the ADH regions. The polynucleotides and
CC polypeptides may also be used for monitoring therapeutic effects during
CC clinical trials and other treatments. The present sequence represents the
CC human 21620 alcohol dehydrogenase protein of the invention the gene which
CC encodes this protein has been mapped to chromosome 17 (17q12-21)
XX
SQ Sequence 260 AA;

```
  Query Match             100.0%;  Score 1337;  DB 5;  Length 260;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-131;
  Matches  260;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGY 60

Qy         61 PGTLIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGTLIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNV 120

Qy        121 NVLALSICTREAYQSMKERNVDDGHIININSMSGHRVLPLSVTHFYSATKYAVTALTEGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NVLALSICTREAYQSMKERNVDDGHIININSMSGHRVLPLSVTHFYSATKYAVTALTEGL 180

Qy        181 RQELREAQTHIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RQELREAQTHIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVL 240

Qy        241 STPAHIQIGDIQMRPTEQVT 260
              ||||||||||||||||||||
Db        241 STPAHIQIGDIQMRPTEQVT 260
```

RESULT 6
ABU58588
ID ABU58588 standard; protein; 260 AA.
XX
AC ABU58588;
XX
DT 15-APR-2003 (first entry)
XX
DE Human PRO polypeptide #189.
XX
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX
OS Homo sapiens.
XX
PN US2003027272-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176492.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.

```
Qy        972 GGATCACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAG 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGATCACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAG 1020

Qy       1032 ATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAA 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAA 1080

Qy       1092 GGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCTGGGCACTTGG 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCTGGGCACTTGG 1140

Qy       1152 CCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCC 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCC 1200

Qy       1212 CATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCAC 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCAC 1260

Qy       1272 CTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCCACTGCACC 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCCACTGCACC 1320

Qy       1332 CTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGGCTTCTTGTCCCCT 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGGCTTCTTGTCCCCT 1380

Qy       1392 CCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGCCC 1451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGCCC 1440

Qy       1452 AGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAAAAAAAAAAAA 1505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAAAAAAAAAAAA 1494
```

RESULT 5
AX179289
LOCUS AX179289 1909 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 2 from Patent WO0144446.
ACCESSION AX179289
VERSION AX179289.1 GI:14598959
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meyers,R.
TITLE 21612, 21615, 21620, 21676, 33756, novel human alcohol
dehydrogenases
JOURNAL Patent: WO 0144446-A 2 21-JUN-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1. .1909
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 421. .1203
/note="unnamed protein product; 21620 ADH"
/codon_start=1
/protein_id="CAC43880.1"
/db_xref="GI:14598960"
/db_xref="REMTREMBL:CAC43880"
/translation="MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCART
VGNIEELAAECKSAGYPGTLIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLAR
PDTLLSGSTSGWKDMFNVNVLALSICTREAYQSMKERNVDDGHIININSMSGHRVLPL
SVTHFYSATKYAVTALTEGLRQELREAQTHIRATCISPGVVETQFAFKLHDKDPEKAA
ATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT"
ORIGIN

Query Match 99.1%; Score 1492; DB 6; Length 1909;

Best Local Similarity 99.8%; Pred. No. 4.5e-276;
Matches 1503; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```
Qy          1 CGCGGATCGGACCCAAGCAGGTCGGCGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CGCGGATCGGACCCAAGCAGGTCGGCGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCC 369

Qy         61 TCGACCCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TCGACCCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGG 429

Qy        121 CCCGGCATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 CCCGGCATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATC 489

Qy        181 GGCGCGGCCGTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 GGCGCGGCCGTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGC 549

Qy        241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACT 609

Qy        301 TTGATCCCCTACAGATGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 TTGATCCCCTACAGATGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 669

Qy        361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCCT 729

Qy        421 GACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG 789

Qy        481 GCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGAAGGAGCGGAATGTGGACGAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 GCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGAAGGAGCGGAATGTGGACGAT 849

Qy        541 GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTACCCCTGTCTGTGACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTACCCCTGTCTGTGACC 909

Qy        601 CACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACTGAGGCAAGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACTGAGGCAAGAG 969

Qy        661 CTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGGAGACA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 CTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGGAGACA 1029

Qy        721 CAATTCGCCTTCAAACTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CAATTCGCCTTCAAACTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA 1089

Qy        781 ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCC 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1090 ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACTCCC 1149

Qy        841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTG 1209

Qy        901 TGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 TGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGG 1269

Qy        961 TGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 TGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAA 1329

Qy       1021 ATTTGTTTGAGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1330 ATTTGTTTGAGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATG 1389

Qy       1081 GGCTGGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCC 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 GGCTGGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCC 1449

Qy       1141 TGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 TGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1509

Qy       1201 GCCAAAATCCCCATCTTCTTGCACCTCAACGTCTGTGGCTCA-GGGCTGGGGTGGCAGAG 1259
             ||||||||||||||||||||||||||||||||||||||||:| ||||||||||||||||||
Db       1510 GCCAAAATCCCCATCTTCTTGCACCTCAACGTCTGTGGCTYANGGGCTGGGGTGGCAGAG 1569

Qy       1260 GGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTG 1319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 GGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTG 1629

Qy       1320 CCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGGC 1379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 CCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGGC 1689

Qy       1380 TTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACC 1439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 TTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACC 1749

Qy       1440 AGGGCCTGGCCCAGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAAAAAA 1499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 AGGGCCTGGCCCAGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAAAAAA 1809

Qy       1500 AAAAAA 1505
              ||||||
Db       1810 AAAAAA 1815
```

RESULT 6
AK026196

```
LOCUS       AK026196               1410 bp    mRNA    linear   PRI 13-SEP-2003
DEFINITION  Homo sapiens cDNA: FLJ22543 fis, clone HSI00212.
ACCESSION   AK026196
VERSION     AK026196.1  GI:10438967
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1410)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .1410
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI00212"
                     /tissue_type="human small intestine"
                     /clone_lib="HSI"
                     /note="cloning vector pME18SFL3"
     CDS             243. .788
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB15390.1"
                     /db_xref="GI:10438968"
                     /translation="MSSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNVL
                     ALSICTREAYQSMKERNVDDGHIININSMSGHRVLPLSVTHFYSATKYAVTALTEGLR
                     QELREAQTHIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYV
                     LSTPAHIQIGDIQMRPTEQVT"
ORIGIN
```

```
  Query Match           83.7%;  Score 1259.8;  DB 9;  Length 1410;
  Best Local Similarity  96.1%;  Pred. No. 1.6e-231;
  Matches 1291;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy        163 GGGGCCTCGGGGGGCATCGGCGCGGCCGTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAG 222
              ||| |  | ||  | | |  |  | |  |||  ||      || | | |
Db         57 GGGTCTAGGCGCGGATCGGACCCAAGCAGGTCGGCGGCGGCGGCAGGAGAGCGGCCGGGC 116

Qy        223 GTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGT 282
              ||      ||  ||| |     ||||||||||||||||||||||||||||||||||||
Db        117 GTCAGCTCCTCGACCCCCGTGTCGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGT 176

Qy        283 GCAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTATCAAATGAAGAGGACATC 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 GCAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTATCAAATGAAGAGGACATC 236

Qy        343 CTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAAT 402
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CTCTCCATGTCCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAAT 296

Qy        403 GCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAAGGACATG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAAGGACATG 356

Qy        463 TTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGAAG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGAAG 416

Qy        523 GAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTG 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 GAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTG 476

Qy        583 TTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 TTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACA 536

Qy        643 GAGGGACTGAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCT 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GAGGGACTGAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCT 596

Qy        703 CCAGGTGTGGTGGAGACACAATTCGCCTTCAAACTCCACGACAAGGACCCTGAGAAGGCA 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 CCAGGTGTGGTGGAGACACAATTCGCCTTCAAACTCCACGACAAGGACCCTGAGAAGGCA 656

Qy        763 GCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATC 716

Qy        823 TACGTCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCACGGAG 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TACGTCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCACGGAG 776

Qy        883 CAGGTGACCTAGTGACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTC 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 CAGGTGACCTAGTGACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTC 836

Qy        943 CTGCCTCTGGATTTTAGGTGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACC 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       122 GCCTCGGGGGGCATCGGTGCGGCCGTGGCCCGGGCATTAGTCCAGCAGGGACTGAAGGTT 181
Qy       226 GTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCA 285
Db       182 GTGGGTTGTGCCCGCACCGTTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCA 241
Qy       286 GGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTATCAAATGAAGAGGACATCCTC 345
Db       242 GGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTGTCAAATGAGGAGGACATCCTC 301
Qy       346 TCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCT 405
Db       302 TCCATGTTCTCAGCTGTCCGATCCCAGCACAGTGGCGTGGATATCTGCATCAACAATGCC 361
Qy       406 GGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAAGGACATGTTC 465
Db       362 GGCATGGCCCGGCCTGACACCCTGCTCTCGGGCAGCACCAGCGGATGGAAGGACATGTTC 421
Qy       466 AATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGAAGGAG 525
Db       422 AATGTGAATGTGCTGGCCCTCAGCATCTGCACTCGGGAGGCTTATCAGTCCATGAAGGAG 481
Qy       526 CGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA 585
Db       482 CGGAACATAGACGACGGGCACATCATTAACATCAACAGCATGTGTGGCCACCGAGTCCCA 541
Qy       586 CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAG 645
Db       542 CCCCAGTCTGTGATCCATTTCTATAGTGCGACTAAGTATGCCGTCACTGCACTGACAGAG 601
Qy       646 GGACTGAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCA 705
Db       602 GGACTCAGGCAAGAGCTTCTGGAGGCCCAGACCCATATCCGGGCCACGTGTATCTCTCCA 661
Qy       706 GGTGTGGTGGAGACACAATTCGCCTTCAAACTCCACGACAAGGACCCTGAGAAGGCAGCT 765
Db       662 GGCTTGGTAGAGACACAGTTCGCCTTCAAACTCCATGACAAGGACCCGGGGGAAGCAGCT 721
Qy       766 GCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTAC 825
Db       722 GCCACCTATGAACACATAAAGTGTCTCAGACCAGAGGACGTGGCTGAGGCTGTCATCTAC 781
Qy       826 GTCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCACGGAGCAG 885
Db       782 GTTCTTAGCACACCCCCACATGTTCAGGTTGGTGACATCCAGATGAGGCCCACAGAGCAG 841
Qy       886 GTGACCTAGTGACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTG 945
Db       842 GTGACCTAGTG-TTCTGTGGGGAGATCCTACCTTCTCCACCCTTCACGGCTCGCCTCCTG 900
Qy       946 CCTCTGGATTTTAGGTGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCG 1005
Db       901 CCTCTGGATTTTAGGTGTTGATTTCTGGA-CTCTGCAGCTTACTTTCTATGTCCACCCCA 959
Qy      1006 ACCAGGGGCTAGAAAATTTGTTTGAGATTTTTATATCATCTTGTCAAATTGCTTCAGTTG 1065
Db       960 CTAAGGG----------------------ATAGACAGCTTGTCAGACTATTCCCATTA 995
Qy      1066 TAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAAC 1125
Db       996 CAAATGTGAAAAACTGGGCCGGGAAGAGAAGACAGAATCTCGACTGTTCTATCTAACTTC 1055
Qy      1126 TTGTTCTTGTGCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACAT 1185
Db      1056 TTCTGAGTCCTCTTGGGGCTCCGTAACCTTTG-CTCTCCTGTGCCTCTTCCTTTTGACCT 1114
Qy      1186 GGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGG 1245
Db      1115 AGGAAAGGGATGTACCTAGAAGGGGCATTGCTCTATCCTCTTACACCCTACCCCCCACCC 1174
Qy      1246 CT--GGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGA 1303
Db      1175 CTTAGGGTTGGAGCAGCAGAGACCTCATCTTTTACCTGTGTGGTTGTCCAGGGTTCCAGG 1234
Qy      1304 CTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCC 1363
Db      1235 TTCCCTTCCCAACCCTCCTGCCTCCCCCCCCCCCAGCTCTCCAGCCAAGTTTTGGCCCTC 1294
Qy      1364 CC 1365
Db      1295 AC 1296
```

```
RESULT 8
AC003042/c
LOCUS       AC003042               102818 bp    DNA     linear   PRI 14-JUL-1998
DEFINITION  Homo sapiens chromosome 17, clone HCIT75G16, complete sequence.
ACCESSION   AC003042
VERSION     AC003042.1  GI:3319121
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 102818)
  AUTHORS   Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone HCIT75G16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 102818)
  AUTHORS   Birren,B., Fasman,K., McKernan,K., Munro,C., Nusbaum,C.,
            Richardson,P., Lander,E., Baldwin,J., Barna,N., Cantu,C., Chang,A.,
            Cooke,P., Daly,M.J., Devon,K., Dewar,K., DuRette,B., Forrest,C.,
            Gage,D., Gensheimer,S., Geraigery,K., Gilmartin,T., Hagos,B.,
            Halphen,I., Harris,K., Howland,J.C., Huang,J., Hui,L., Jacotot,L.,
            Kirby,A., Lane,M., MacKenzie,J., Marquis,N., McDermott,J.,
            Molla,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A.,
            Peterson,K., Roberts,D., Rollins,G., Sarnaik,A., Shiu,P., Shyam,R.,
            Stilwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L.,
            Zemtseva,I. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-1997) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 102818)
  AUTHORS   Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatin,C.,
            Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
            Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
            FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
            Gensheimer,S., Geraigery,K., Gilmartin,T., Grant,G., Hagos,B.,
            Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
            Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
            Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
            Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
            Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-1998) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 14, 1998 this sequence version replaced gi:3294535.
            All repeats were identified using RepeatMasker: Smit, A.F.A. &
            Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES             Location/Qualifiers
     source          1. .102818
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="HCIT75G16"
                     /clone_lib="Research Genetics/Cal Tech CITB978SK-B (plates
                     1-194)"
```

```
repeat_region   495. .778
                /rpt_family="AluJb"
repeat_region   complement(1357. .1409)
                /rpt_family="AT_rich"
repeat_region   1737. .1800
                /rpt_family="(CA)n"
repeat_region   1911. .1964
                /rpt_family="MIR"
repeat_region   2074. .2216
                /rpt_family="MIR"
repeat_region   complement(2965. .2985)
                /rpt_family="AT_rich"
repeat_region   3225. .3255
                /rpt_family="(GA)n"
repeat_region   3352. .3446
                /rpt_family="(GGA)n"
repeat_region   complement(3681. .3713)
                /rpt_family="AT_rich"
repeat_region   complement(3966. .4112)
                /rpt_family="MIR"
repeat_region   complement(4970. .5069)
                /rpt_family="L2"
repeat_region   5648. .5772
                /rpt_family="MIR"
repeat_region   5957. .5996
                /rpt_family="MIR"
repeat_region   complement(6226. .6270)
                /rpt_family="(TGG)n"
repeat_region   6890. .6916
                /rpt_family="GC_rich"
repeat_region   7990. .8224
                /rpt_family="AluSx"
repeat_region   8230. .8737
                /rpt_family="L1MB8"
repeat_region   8742. .9020
                /rpt_family="L1MB8"
repeat_region   complement(9027. .9152)
                /rpt_family="MER5A"
repeat_region   10394. .10459
                /rpt_family="L2"
repeat_region   complement(10478. .10504)
                /rpt_family="(CAT)n"
repeat_region   complement(10582. .10881)
                /rpt_family="AluSg"
repeat_region   complement(10942. .11013)
                /rpt_family="MIR"
repeat_region   complement(11900. .11931)
                /rpt_family="MIR"
repeat_region   12398. .12694
                /rpt_family="AluJb"
repeat_region   complement(12919. .13084)
                /rpt_family="FRAM"
repeat_region   14235. .14354
                /rpt_family="L1MC/D"
repeat_region   15043. .15089
                /rpt_family="(GA)n"
repeat_region   complement(15240. .15355)
                /rpt_family="(GGA)n"
repeat_region   complement(15360. .15426)
                /rpt_family="(GGA)n"
repeat_region   16790. .16915
                /rpt_family="AluSq/x"
repeat_region   complement(17391. .17454)
                /rpt_family="MIR"
repeat_region   complement(17689. .17749)
                /rpt_family="MIR"
repeat_region   complement(17800. .17927)
                /rpt_family="MIR"
repeat_region   18030. .18086
                /rpt_family="L2"
repeat_region   complement(18850. .18885)
                /rpt_family="MIR"
repeat_region   complement(19951. .20001)
                /rpt_family="(GA)n"
repeat_region   20054. .20157
                /rpt_family="MIR"
repeat_region   20635. .20762
                /rpt_family="FLAM_C"
repeat_region   20947. .21243
                /rpt_family="AluSx"
repeat_region   22098. .22153
                /rpt_family="MIR"
repeat_region   22268. .22558
                /rpt_family="AluSx"
repeat_region   complement(22631. .22688)
                /rpt_family="TIGGER2"
repeat_region   23162. .23460
                /rpt_family="AluSx"
repeat_region   23461. .23512
                /rpt_family="(GGA)n"
repeat_region   23515. .23637
                /rpt_family="(GAA)n"
repeat_region   complement(23681. .23761)
                /rpt_family="L2"
repeat_region   complement(26353. .26652)
                /rpt_family="AluSx"
repeat_region   26941. .27047
                /rpt_family="(CAGA)n"
repeat_region   complement(27894. .28005)
                /rpt_family="MIR"
repeat_region   28673. .29257
                /rpt_family="L1MB4"
repeat_region   complement(29262. .29556)
                /rpt_family="AluSx"
repeat_region   29562. .30018
                /rpt_family="L1MB4"
repeat_region   30035. .30160
                /rpt_family="MIR"
repeat_region   30320. .30669
                /rpt_family="L2"
repeat_region   31247. .31550
                /rpt_family="AluSx"
repeat_region   complement(31605. .31741)
                /rpt_family="MIR"
repeat_region   31919. .32211
                /rpt_family="AluJo"
repeat_region   complement(32934. .33227)
                /rpt_family="AluSx"
repeat_region   complement(33268. .33560)
                /rpt_family="AluSp"
repeat_region   complement(34210. .34319)
                /rpt_family="MIR"
repeat_region   34835. .34869
                /rpt_family="(CAT)n"
repeat_region   complement(34876. .35054)
                /rpt_family="MIR"
repeat_region   complement(36108. .36327)
                /rpt_family="MIR"
repeat_region   complement(36957. .37440)
                /rpt_family="MLT2CB"
repeat_region   37827. .38129
                /rpt_family="AluSx"
repeat_region   complement(38139. .38249)
                /rpt_family="MER5B"
repeat_region   38410. .38708
                /rpt_family="AluY"
repeat_region   38730. .39009
                /rpt_family="AluSx"
repeat_region   complement(39165. .39287)
                /rpt_family="MIR"

Query Match             42.9%;  Score 646;  DB 9;  Length 102818;
Best Local Similarity   100.0%;  Pred. No. 8.2e-114;
Matches  646;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        850 CAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTC 909
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48792 CAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTC 48733

Qy        910 CTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTT 969
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48732 CTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTT 48673

Qy        970 CTGGATCACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTG 1029
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48672 CTGGATCACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTG 48613

Qy       1030 AGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGA 1089
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48612 AGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGA 48553

Qy       1090 AAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCTGGGCACTT 1149
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48552 AAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCTGGGCACTT 48493

Qy       1150 GGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATC 1209
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48492 GGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATC 48433

Qy       1210 CCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTC 1269
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48432 CCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTC 48373

Qy       1270 ACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCCACTGCA 1329
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48372 ACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCCACTGCA 48313

Qy       1330 CCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGGCTTCTTGTCCC 1389
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48312 CCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGGCTTCTTGTCCC 48253

Qy       1390 CTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC 1449
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48252 CTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC 48193

Qy       1450 CCAGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA 1495
                |||||||||||||||||||||||||||||||||||||||||||||
Db      48192 CCAGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA 48147
```

RESULT 9
AC110594
LOCUS AC110594 159490 bp DNA linear PRI 13-SEP-2002
DEFINITION Homo sapiens chromosome 17, clone CTD-2193J24, complete sequence.
ACCESSION AC110594
VERSION AC110594.5 GI:22830329
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159490)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone CTD-2193J24
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 159490)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 159490)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 159490)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Sep 13, 2002 this sequence version replaced gi:22296745.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L25381
Center clone name: 2193_J_24

Only the first 159.5 kilobases of this clone are being submitted. The remainder overlaps accession number AC003042 [WICGR project L260].